**Identifying Important Identities: What Individuals Value in Themselves and Others**

Isaiah R. Raynal

[raynal2@illinois.edu](mailto:raynal2@illinois.edu) | ORCID: 0000-0003-2852-4650

LinkedIn: <https://www.linkedin.com/in/isaiahraynal/>

Website: <https://www.isaiahraynal.com/>

Department of Political Science, University of Illinois Urbana-Champaign

**Abstract**

**Keywords**

conceptualization and measurement, self-identification, social identity, relationships, open-ended survey, content analysis, United States

**Introduction**

Academic researchers across social science disciplines are increasingly focused on identity in their research. There are empirical articles using it as both an explanatory and sometimes as an outcome variable to better understand people’s individual and social identities. There is also a literature on the conceptualization and measurement of identity. In this article, I argue that there are limitations to these previous discussions of identity in that substantively scholars are often hyper-focused on only one type of identity (i.e. demographics) and are often limited in their definitions of the term (i.e. academic ones). To address this gap, I present the results of an original open-ended survey that directly asked ordinary Americans what identities are most important to them for themselves and in their relationships to better understand the wide variety of ways in which people can identify.

Identities are important to understand because they have many implications for people’s behavior. People want to take action that is consistent with their identity, which is known as “identity-based motivation” (Oyserman et al., 2007) or “identity congruence” (Shang et al., 2008). Identities can affect behaviors such as donations (Shang et al., 2008; Aaker and Akutsu, 2009) as well as voting (Kalin and Sambanis, 2018). By better understanding how people identify, we can better understand how they behave and why.

**Research Questions**

What identities do people value most in themselves and others (EQ1), both real (EQ2) and hypothetical (EQ3)? My research question is largely exploratory in that I am interested in learning what identities are most important to people; I do not have a priori expectations regarding which particular identities will emerge as most notable from this inquiry. To the contrary, my one expectation going into this study is that there will be a variety in the identities that people feel are important to them.

The way I talk about identity in this chapter may differ from the way that other scholars discuss it. This is because while there may be certain academic definitions of identity, I am most interested in how people identify themselves and others, which involves a more colloquial understanding of the term. Dictionary.com, for example, defines identities as “all kinds of characteristics, qualities, experiences, interests, and other aspects of a person.” Part of the purpose of this chapter is to understand how Americans see themselves and others, and I find it unlikely that the average person (i.e. non-academic) will differentiate between identities, traits, personality type, etc. when thinking about their concept of self. As a result, I will be focusing on attributes as a whole and will use the terms “identity” and “attributes” interchangeably.

I also purposely chose to ask Americans directly which identities are most important to them because I do not want to make any false assumptions. Closed-ended survey questions are often limited precisely because of this reason: they are pre-determined assumptions made by the researcher (Hutsebaut and Verhoeven, 1995; Reja et al., 2003) that respondents must select from even if those are not the options they would have chosen in the first place. One solution to soliciting responses when the responses may be unknown is to ask open-ended questions (Züll, 2016). I administered a survey that did precisely that. In this article, I discuss the design and findings of an open-ended survey that asked respondents how they identify, which of these identities they share with the relationships in their lives, and which they care about in potential future relationships.

The rest of this article proceeds as follows. First, I take a deeper dive into previous literature on identity. I show that social science scholars have typically focused on certain definitions of identity (i.e. academic ones) as well as certain types of identities (i.e. demographic ones). I then advocate for my methodological and substantive contributions of asking ordinary people how they identify to better learn the different ways people understand identity. I also make the case for why I administered an open-ended survey in my pursuit of these goals. Second, I describe a survey that I administered to 233 American respondents asking them which identities are most important to them. Third, I present results from a quantitative content analysis and qualitative coding, both done by hand, as well as some computational text analyses. I find that the most common type of identity mentioned are those related to personalities, behaviors, and traits, specifically “honest,” “friend,” “fun,” “loyal,” and “trust.” Finally, I conclude with a summary of the main findings as well as the importance of this research for a variety of disciplines interested in the study of identity.

**The Multiplicity of Individual and Social Identities as Understood by Ordinary Americans**

Scholars across social science are increasingly interested in identity both as an explanatory and an outcome variable in their research. However, these academic researchers often make a variety of assumptions about how ordinary people 1) self-identify and 2) understand the term “identity” in the first place. This article makes both a substantive and methodological contribution to the interdisciplinary literature on identity by exploring both of the aforementioned assumptions, respectively.

In terms of how people self-identify, researchers and other academics often focus on demographic identities. There are entire departments and journals dedicated to better understanding demographic identities related to race, ethnicity, gender, sexuality, and more. Literature focused on intergroup relations often use demographic identities in their treatments to promote contact and positive affect between outgroup members, such as Levendusky (2018) did with national identity and Klar (2018) tried with gender in attempts to reduce affective polarization between Republicans and Democrats. While demographic identities are nonetheless important, scholars often narrow focus on these types of identities could be missing other ways in which people identify. For example, Sen (2006) discusses how people identify with a variety of labels, both those related to demographics such as nationality, religion, and gender as well as other categories such as occupation, hobby, and dietary restrictions.

Literature on the conceptualization and measurement of identity are often focused on scholarly definitions of the term (e.g. Fearon, 1999; Brubaker and Cooper, 2000). However, the concept of “identity” is prevalent in the lives of ordinary people, who may have understandings of the term different from the academy.

To address both these limitations, I administer an original open-ended survey asking people which identities they most identify with as well as which ones they share with the people in their lives or wish they shared with future connections. Open-ended questions are an effective tool for being able to hear directly from respondents. In open-ended responses, respondents can say whatever they are thinking, providing a more diverse set of answers, in contrast to closed-ended response options, which are often limited by the researcher who designed them’s assumptions (Hutsebaut and Verhoeven, 1995; Reja et al., 2003). The problem with having respondents select from a premade list of identities (e.g. the National Election Study’s “group closeness” question, see Wong, 1998) is that someone’s most important identities may or may not be on that list

Open-ended surveys are an effective way of asking people about their identities specifically as they can choose how to describe themselves instead of having to pick from a list (McGuire and Padawer-Singer, 1976). Symbolic interactionist theory claims that the best definitions of a person’s identity are those made by one’s own self (Spitzer et al., 1971). Identity scholars have long used a variety of open-ended questions, such as the “Who Are You?” (W-A-Y) test (Bugental and Zelen, 1950) and the Twenty Statements Test (TST) (e.g. Kuhn and McPartland, 1954; Spitzer et al., 1971), which is considered “a standardized measure of self-attitudes” (Kuhn and McPartland, 1954). These tests have been used to “allow the client to structure his responses along lines most expressive of his own needs and most meaningfully related to his current situation” (Bugental and Zelen, 1950 p. 484). Scholars in these disciplines who have used open-ended questions such as those I just mentioned have found a variety of other identities across many categories such as demographics, likes and dislikes, skills, personal values, goals, and more (e.g. Kuhn and McPartland, 1954; Spitzer et al., 1971; Eccles, 2009). This is similar to the goal of my project that seeks to identify a wide variety of attributes instead of limiting the number of possibilities with closed-ended questions.

My survey begins with an open-ended question follower by several closed-ended questions in which the survey is programmed to show respondents what they wrote earlier and then to select which of those apply to various follow-up questions. The open-ended question asks respondents to make a list of different aspects of themselves and others that they think are important and then follow-up closed-ended questions ask which of those they share with those in their lives or which are most important in hypothetical future relationships, depending on the version of the survey they receive. I want the initial open-ended question in order to learn which identities are most important to people free from any assumptions. But, since the follow-up questions involve choosing from that list, I program the survey to make sure they are in fact selecting from their initial list.

**Materials and Methods**

**Survey Design**

I administered a descriptive survey to 233[[1]](#footnote-1) respondents in July 2023 using Lucid Theorem. The survey was funded by the Department of Political Science at the University of Illinois Urbana-Champaign. The survey instrument can be found in Appendix A. Respondents self-selected in for pay ($1.50), however, Lucid claims that the distribution of their demographics matches those of a nationally representative sample.Furthermore, researchers such as Krupnikov et al. (2021) have shown that results from convenience samples tend to reliably replicate to true nationally representative samples. Respondents were randomly assigned to one of five groups. The randomization was not for experimental purposes but rather to combine multiple surveys into one.

Respondents in the first two groups were asked an open-ended question about which identities[[2]](#footnote-2) are most important to them (EQ1) followed by closed-ended questions asking respondents which of those identities they share with the people in their life (EQ2). One group received questions that asked generically about the people in their lives while the other group received questions that asked them to think about specific people in their lives. Both groups were asked about family members, friends, acquaintances, and people whom they dislike. The purpose of the questions overall was to identify respondents’ own most important identities as well as which similar or dissimilar identities come up in respondents’ relationships. I purposely asked about a variety of relationships (close versus weak ties, liked versus disliked people) to see whether shared identities change depending on the type of relationship. This is because there are many different types of relationships and the context of those relationships matter for levels of intimacy (Altman and Taylor, 1973), and I expect people to share more identities with whom they share higher levels of intimacy.

The other three groups were asked to imagine a hypothetical scenario in which they have some choice over a relationship in their life and were asked which shared identities they would like to prioritize in that relationship (EQ3). The purpose of these groups was to vary the context and see whether people prioritize identities differently in different contexts. Again, this was not for experimental purposes, but rather to combine different versions of a survey in one for economic purposes as opposed to administering multiple different surveys. However, the results will only be compared descriptively, and I will be making no causal arguments about the question wording. The three contexts to which a respondent could be assigned were a friend, a neighbor, or a coworker. In all contexts, respondents were asked to list the top 10 identities, attributes, or characteristics they care most about in these relationships, to narrow down those top 10 identities to the top 3, and to pick which is the most important.

The survey asked about real as well as hypothetical relationships because both are important in different ways. On one hand, we want to understand and know about people’s identities and relationships in the real world. On the other hand, people may have preferences that are constrained by their circumstances; for example, someone who is a member of a racial minority group may want more friends of their own race, but - if they are in a predominantly white area - they may have mostly white friends due to who is available in their neighborhood. Therefore, the questions about hypothetical relationships will get at what their desired relationships look like, which may or may not differ from their actual relationships.

**Survey Sample**

A total of 233 respondents answered the survey. Respondents ranged from age 18 to 94, with the median age being 43. There were 133 women and 100 men - Lucid Theorem only provided responses for those two genders. 78.0 percent of respondents were white. 45.5 percent of respondents identified as Democrats or said they leaned Democratic, 32.6 percent identified as Republicans or said they leaned Republican, and 21.9 percent identified with neither or with another political party.

The 233 respondents were distributed relatively evenly[[3]](#footnote-3) across the five groups.

*[Table 1]*

Between 40 and 50 respondents answered each version of the survey. Each respondent listed between 1 and 10 identities, which resulted in 1,446 responses. After removing responses that were equivalent to being left blank, such as putting “don’t know” or “nothing,” as well as random words such as “I’m,” I was left with 1,432 responses, which is the N I will be using for my analyses.

**Methods**

There are two main analyses I performed and will share in this chapter. First, I had two undergraduate research assistants perform hand-coded content analysis of the open-ended text data in order to categorize the responses into different types of identities, characteristics, and traits. Second, I used computational analyses to identify the most commonly used words across responses. The content analysis was broader in being able to combine characteristics into larger categories, such as demographics, hobbies, or personality types, while the computational text mining was used to identify specific words being used such as honest or kind.

***Content Analysis***

First, I had two undergraduate research assistants perform a content analysis on the open-ended text data in order to group responses into larger categories in order to learn what types of identities are most important to people. I took an iterative approach to content analysis, using Neuendorf (2017) and Tracy (2020) as guides. I wrote an initial codebook with categories I expected to see in the data as well as ones I saw after briefly skimming the responses. I then trained two undergraduate research assistants in content analysis, and they performed several rounds of practice coding on random samples of the data in order to revise the codebook. The final version of our codebook - which can be found in Appendix B - contained the following categories along with some examples of each:

* Demographics: Age, gender, sexuality, race and ethnicity, education level, religion, family role, occupation. Responses were sometimes specific (ex. “male,” “bisexual,” “Irish”) and sometimes general (ex. “race,” “gender,” “age”)
* Political views (ex. “conservative,” “progressive”)
* Hobbies and interests (ex. “dancing,” “nature,” “art,” “musician”)
* Physical body (ex. “balding,” “healthy,” “pretty”)
* Personality, behaviors, and other traits (ex. “honest,” “friendly,” “hardworking”)
* None of the above

The unit of analysis was the response, and the undergraduates coded which of the above variables was present in the response. Coders were told to code only each response as exactly one variable; responses could not be coded as more than one variable, nor could they be left blank.

I calculated one measure of intercoder reliability by creating one column per coder that contained which variable they coded each response as. In other words, there was one column with Coder 1’s name and another column with Coder’s 2 name, and the values within each column were what the coder labeled each response as. As an example, if both coders labeled the attribute as a demographic, that was understood as an agreement. In contrast, if one coder labeled an attribute as a demographic but the other coder labeled it as none of the above, then that was marked as a disagreement.

The table below shows the intercoder agreement statistics: both raw agreement (percent agreement) as well as chance-corrected agreement (Cohen’s Kappa and Krippendorff’s Alpha). The statistics were calculated using Dr. Deen Freelon’s online resource ReCal (“reliability calculator”).

*[Table 2]*

There are a variety of perspective on cutoffs for reliability, but Neuendorf (2017)’s review of the work on reliability comes to the conclusion that, for chance-corrected agreement (such as Cohen’s kappa), 0.80 or higher would be acceptable to all and between 0.6 and 0.8 would be acceptable to most. Therefore, the relatively high statistic of 0.787 should be considered acceptable.

***Text Mining***

Second, I employed computational methods in order to identify the most commonly used words in order to discover the specific types of identities that were most common across responses. For example, to foreshadow the results, the content analysis revealed that personality, behavior, and traits were the most common types of identity mentioned, while the computational text mining revealed that “honest” was the most commonly mentioned specific identity.

I first unnested the responses into single words, changing each row to be by single word instead of by response in order to allow me to later combine repeated words and then count the frequency of each unique word. I then tried to computationally stem and combine similar words (such as “honest,” “honesty,” “honestly,” etc.) however there were two problems with this: the code was 1) missing certain words (ex. “acceptance” and “accepting” being both under “accept”) and 2) combining some words that are different even though they sound the same (ex. “person” and “personality” into “person”). As a result, with the help of an undergraduate assistant, we manually went through all words, identified which words should be combined, and wrote the code to combine them ourselves. This also allowed us to identify and correct spelling errors, such as “haapy” instead of “happy.” We then calculated the word frequencies for all unique words in order to see which appeared most frequently.

**Results**

To reiterate, I am interested in 1) the identities people most frequently self-identify with and share with those with whom they have relationships and 2) the identities people care about most in others. At the end of this section, I will also show whether these results differ by partisanship.

**What Identities do People Value for Themselves?**

In two versions of the survey, respondents were asked to list up to top 10 identities, characteristics, or traits with which they identify. Figure 1 below shows the results from the content analysis categorizing their responses.

*[Figure 1]*

As can be seen in Figure 1, respondents overwhelmingly self-identified with personality, behaviors, and traits (N = 254) the most. Demographics (N = 44) were the second most common followed by hobbies & interests (N = 20) then physical characteristics (N = 18). Only 1 respondent assigned to one of the survey versions about their own identities listed a political view.

After being asked to list their top ten identities, respondents were asked to select which of the identities they listed were shared with the people in their lives. Table 3 below shows, out of the number of responses for each category, what percentage of each of them were shared with family members, friends, acquaintances, and people they dislike.

*[Table 3]*

People shared the most identities with their family members and with friends, both of which are close ties.[[4]](#footnote-4) The main difference is that people are more likely to share demographic identities with family members than with friends, which makes sense, as people are typically biologically related to - and therefore share certain demographic identities such as race and ethnicity with - their family members. When it comes to weaker ties, people share fewer identities with acquaintances and even fewer with people they dislike. These results seem intuitive in that people are more likely to have relationships with those with whom they are similar. These results also remained similar even when split by survey version in terms of whether people were asked to think about relationships in general or a specific person with whom fit into that relationship category; the lack of notable differences between the general and specific groups suggest that thinking about relationships in general versus about a specific person did not affect people’s answers.

*[Table 4]*

*[Table 5]*

**What Identities do People Care About in Others?**

The previous section showed which identities people share with the people in their lives, which is important to know in order to understand ingroup and outgroup social identities in reality. However, people do not always have full control over the people with whom they have relationships. Family is, more often than not, that into which we are born. Friends are typically thought of as chosen, but they are also often constrained by certain factors such as geography; for example, you are more likely to have friends in the area in which you live because you are more likely to meet people in the area in which you live, whether you like the area in which you live and its inhabitants are not always relevant. Acquaintances and people you dislike are similarly constrained in terms of being people you bump into at the grocery store, friends of friends, coworkers, etc.; you might be able to control the frequency with which you see them, but not whether you see them at all. All of this is to say that the identities people share with those in their lives may or may not be the identities they wish to share. For example, a person of color living in a predominantly white neighborhood might wish to have friends of their own race, but do not because of where they live. Similarly, a white person living in a predominantly white neighborhood may wish to have friends of different backgrounds but cannot do so due to their homogeneous community. Therefore, the other versions of the survey asking about hypothetical relationships aim to understand what types of identities people would like to have in their relationships. Since the type of relationship could matter, I asked about three different ones - coworker, friend, and neighbor - in order to see what people said across different contexts, all of which are common relationships that almost every person has.

In this section, respondents were first asked to list up to 10 identities, characteristics, and traits that they care about in the relationship about which they were asked.

*[Table 6]*

As can be seen in Table 6 above, respondents cared most about the personality, behaviors, and traits of their hypothetical coworkers, friends, and neighbors. This makes sense given the findings of the other survey versions: if people care about personality, behaviors, and traits for themselves and they share those identities with others, they are also more likely to care about those in general. One difference we can see is that respondents cared more about the demographics of their neighbors than they did about their coworkers and friends.

Respondents were then asked to choose 3 identities they cared most about out of the 10 that they listed. While frequency of mentioned identities may imply importance, asking people directly gets a stronger measure of what someone finds important. For example, someone could have listed nine personality traits and one demographic trait but care the most about the demographic trait. However, Table 7 below shows that the results from the top 3 selected were similar to the top 10 originally listed: overall, people cared most about personality, behaviors, and traits in all relationships and people cared more about the demographics of their neighbor than they did of their coworkers and friends.

*[Table 7]*

Finally, respondents were asked to choose the identity they prioritized most in these relationships, and the results from this question can be found in Table 8 below.

*[Table 8]*

The results from this question are consistent with those of the previous question, showing that respondents do really care most about personalities, behaviors, and traits.

This leaves us with the question of what types of personalities, behaviors, and traits respondents care most about. Table 9 below shows the words mentioned most frequently in the top 10 responses provided by people asked about their own identities. Remember that these include the stems of words, so “friend” would include “friendly,” “friendliness,” “friends,” and “friendship.”

*[Table 9]*

The results of this table are consistent with what we found in the content analysis in terms of being about people’s personalities, behaviors, and traits. Here, we can see more specific examples of the types of characteristics people identify with the most: honest, love, fun, loyal, and trust.

Table 10 below shows the top 10 most common words - top 11 for neighbor, since there were two words tied for 10th place - across responses for desired identities in the three different relationship types.

*[Table 10]*

The results shown in Table 10 are similar to those found in the other versions of the survey, with some slight differences across contexts. The results in the table above help us to see differences in the specific types of identities, traits, and characteristics that people care about in different relationships. We can see some differences that make sense based on the type of relationship being asked about. People want coworkers who are hardworking, smart, and reliable; friends who are loyal and friendly; and neighbors who are quiet. Of course, certain positive traits - such as being honest, trustworthy, and fun - are desirable in personal relationships in general, so it makes sense they would be popular across survey versions.

**Discussion**

This chapter provided an analysis of an open-ended survey that asked Americans which identities are most important to them (EQ1), which of those identities they share with the people in their lives (EQ2), and which identities they would like to share in future hypothetical relationships (EQ3). Across the board, respondents provided answers that fell under the personality, behavior, and traits category such as “honest,” “friend,” “fun,” “loyal,” and “trust.” These results remained consistent across survey versions with slight differences in the most common words across contexts, such as wanting coworkers to be hardworking and neighbors to be quiet. Respondents were more likely to share identities with close ties (family and friends) than weaker ties (acquaintances) and less so with people they dislike. While the results of the content analysis showed personality, behaviors, and traits to be the most common category, it also showed that there were a wide variety of identities that people listed with varying frequency. This variance in identities mentioned supports Sen (2006)’s claims of the multiplicity of people’s identity.

Research on identity often focuses on one particular type of identity: political scientists focus on political identity, racial and ethnic scholars focus on racial and ethnic identity, religious scholars focus on religious identity, and so on. This work is important in showing both how people identify with a variety of labels and how they prioritize and rank those identities. Scholars interested in identity can use these results to inform future work. For example, one political scientist used these results to design two survey experiments testing whether encouraging members of opposite political parties to think about the other identities found in this study can reduce animosity between them (Raynal, 2025). The interventions were successful, perhaps because identities such as positive characteristics are zero-sum and less likely to face the same backlash that heated demographic identities can.

While this research is novel and valuable, it nonetheless has some weaknesses to address. One is in the data: not all respondents filled out all 10 responses. If I were to re-run this study, I would use a platform such as CloudResearch connect where respondents are more motivated to respond to all questions and where I can reject participants who do not provide responses to all the questions. Another limitation is in the analysis: personality, behaviors, and traits is a broad category, and its broadness may be one reason why it is the largest. Originally, the codebook had separate variables for categories such as positive characteristics and personality traits, but practice coding showed we could not separate them with high intercoder agreement as it was difficult to come up with definitions separating the two. For example, is “friendly” a positive characteristic (since it is typically considered to be normatively desirable) or a personality trait (since it is a feature of extroversion)? It therefore made sense to combine them into a larger category that covered descriptors of people’s personality and valent traits.

Another potential limitation is the priming of the definition by Dictionary.com provided at the beginning of the survey. My rationale for including this definition was to give respondents some idea of what I meant by identity without priming them to think about specific types of identities. However, the way that Dictionary.com defined identity may still have influenced the way respondents conceptualized that term. Future research could repeat this study with different prompts that included no definitions, different definitions, or gave a list of examples to see how the different prompts influence results.

Despite these potential limitations, these results are informative for a variety of disciplines that study identity, such as political science, sociology, psychology, and more. The main contribution of these results to the broader literature on identity is that people think of identities as including traits such as honesty and loyalty, and these are the types of identities they identify with the most and prioritize the most in their relationships.

**Acknowledgments**

This article was adapted from a portion of my doctoral dissertation. Thank you to my committee members – Jeffery Mondak, Matthew Winters, Cara Wong, and JungHwan Yang – for your constructive feedback. And thank you to Isabel Escapa and Sam Ya for your help with the analysis; thank you to the Undergraduate Research Apprenticeship Program at the University of Illinois Urbana-Champaign for connecting the three of us and for giving them the opportunity to present at the Undergraduate Research Symposium.

**Declaration of Interest Statement**

No potential conflict of interest was reported by the author.

**Funding**

The author disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study has received funding from the Dean S. Dorman Scholarship from the Department of Political Science at the University of Illinois Urbana-Champaign.

**Data Availability Statement**

**References**

Aaker, J. L., & Akutsu, S. (2009). Why do people give? The role of identity in giving. *Journal of Consumer Psychology, 19,* 267-270.

Altman, I., & Taylor, D. A. (1973). *Social penetration.* New York: Holt, Rinehart & Winston.

Bugental, J. F. T., & Zelen, S. L. (1950). Investigations into the self-concept—the W-A-Y technique. *Journal of Personality, 18,* 483-498.

Dictionary.com. (n.d.). identity. Dictionary.com. <https://www.dictionary.com/browse/identity>

Eccles, J. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist, 44*(2), 78-89.

Freelon, D. (n.d.). Recalfront: A tool for recalculating survey weights for comparative analysis. Retrieved from <https://dfreelon.org/utils/recalfront/>.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, *78*, 1360-1380.

Hutsebaut, D., & Verhoeven, D. (1995). Studying dimensions of God representation: Choosing closed or open-ended research questions. *International Journal for the Psychology of Religion, 5*(1), 49-60.

Kalin, M., & Sambanis, N. (2018). How to think about social identity. *Annual Review of Political Science*, *21*, 239-257.

Klar, S. (2018). When common identities decrease trust: An experimental study of partisan women. *American Journal of Political Science, 62*(3), 610-622.

Krupnikov, Yanna, H. Hannah Nam, and Hillary Style. (2021). Convenience samples in political science experiments. In *Advances in experimental political science,* eds. James N. Druckman and Donald P. Green. New York: Cambridge University Press, 165-83.

Kuhn, M. H., & McPartland, T. S. (1954). An empirical investigation of self attitudes. *American Sociological Review, 19,* 68-76.

Levendusky, M. S. (2018). Americans, not partisans: Can priming American national identity reduce affective polarization? *Journal of Politics, 80*(1), 58-70.

McGuire, W. J., & Padawer-Singer, A. (1976). Trust salience in the spontaneous self-concept. *Journal of Personality and Social Psychology, 33*(6), 743-754.

Neuendorf, K.A. (2017). *The content analysis guidebook* (2nd ed.). SAGE Publications, Inc.

Oyserman, D., Fryberg, S., & Yoder, N. (2007). Identity-based motivation and health. *Journal of Personality and Social Psychology, 93*, 1011−1027.

Raynal, I. R. (2025). *People are more than their politics: Exploring our sense of self as a way to alleviate affective polarization.* (Doctoral dissertation, University of Illinois Urbana-Champaign).

Reja, U., Manfreda, K. L., Hlebec, V., & Vehovar, V. (2003). Open-ended vs. close-ended questions in Web questionnaires. *Developments in Applied Statistics*, *19*, 159-177.

Shang, J., Reed, A., & Croson, R. (2008). Identity congruency effects on donations. *Journal of Marketing Research, 45*(3), 351-361.

Spitzer, S., Couch, C., & Stratton, J. (1971). *The assessment of self.* Iowa City: Escort-Sernole.

Tracy, S.J. (2020). *Qualitative research methods: Collecting evidence, crafting analysis, and communicating impact* (2nd ed.)*.* Wiley Blackwell.

Wong, C. (1998). *Group closeness*. National Election Study Pilot Report. <https://electionstudies.org/wp-content/uploads/2018/04/nes008587.pdf>

Züll, C. (2016). Open-ended questions. *GESIS Survey Guidelines.*

**Appendix A: Survey Instrument**

|  |  |  |
| --- | --- | --- |
| **Question Topic** | **General Question** | **Specific Question** |
| Individual Identities | List the top 10 identities most important to you. Put another way, list the top 10 aspects about yourself with which you identify the most. According to dictionary.com, an identity can include “all kinds of characteristics, qualities, experiences, interests, and other aspects of a person.” (The order in which you list the identities does not matter). | |
| Family Member | Think about your **closest family members**. Of the identities you mentioned earlier, please check the ones you share with those people and leave blank all the identities that you do not share with them.    Identity 1  Identity 2  Identity 3  Identity 4  Identity 5  Identity 6  Identity 7  Identity 8  Identity 9  Identity 10 | Now we’d like to ask you about one of your **close family members**. Please think about someone who is a very close family member of yours and type that person’s initials here: \_AB\_    Earlier, you listed your top identities. Please check all of the identities that you share with \_AB\_ and leave blank all the identities that you do not share with them.    Identity 1  Identity 2  Identity 3  Identity 4  Identity 5  Identity 6  Identity 7  Identity 8  Identity 9  Identity 10 |
| Close Friend | Think about your **closest friends**. Of the identities you mentioned earlier, please check the ones you share with those people and leave blank all the identities that you do not share with them.    Identity 1  Identity 2  Identity 3  Identity 4  Identity 5  Identity 6  Identity 7  Identity 8  Identity 9  Identity 10 | Now we’d like to ask you about one of your **close friends**. Please think about someone who is a very close family member of yours and type that person’s initials here: \_CD\_    Earlier, you listed your top identities. Please check all of the identities that you share with \_CD\_ and leave blank all the identities that you do not share with them.    Identity 1  Identity 2  Identity 3  Identity 4  Identity 5  Identity 6  Identity 7  Identity 8  Identity 9  Identity 10 |
| Acquaintance | Now think about your **acquaintances** such as neighbors or coworkers that you see on occasion but do not know very well. Of the identities you mentioned earlier, please check the ones you share with those people and leave blank all the identities that you do not share with them.    Identity 1  Identity 2  Identity 3  Identity 4  Identity 5  Identity 6  Identity 7  Identity 8  Identity 9  Identity 10 | Now we’d like to ask you about one of your **acquaintances**. Please think about someone who you see on occasion but do not know very well and type that person’s initials here: \_EF\_    Earlier, you listed your top identities. Please check all of the identities that you share with \_EF\_ and leave blank all the identities that you do not share with them.    Identity 1  Identity 2  Identity 3  Identity 4  Identity 5  Identity 6  Identity 7  Identity 8  Identity 9  Identity 10 |
| Disliked Person | Finally, think about **people whom you dislike**. These can be people you know personally or just know of. Of the identities you mentioned earlier, please check the ones you share with those people and leave blank all the identities that you do not share with them.    Identity 1  Identity 2  Identity 3  Identity 4  Identity 5  Identity 6  Identity 7  Identity 8  Identity 9  Identity 10 | Now we’d like to ask you about **someone you dislike**. Please think about someone who you dislike and type that person’s initials here: \_GH\_    Earlier, you listed your top identities. Please check all of the identities that you share with \_GH\_ and leave blank all the identities that you do not share with them.    Identity 1  Identity 2  Identity 3  Identity 4  Identity 5  Identity 6  Identity 7  Identity 8  Identity 9  Identity 10 |

|  |  |  |
| --- | --- | --- |
| **Context 1: Friend** | **Context 2: Neighbor** | **Context 3: Coworker** |
| Imagine you recently moved to a new city and want to start making new friends. You already have one friend in the city who has offered to introduce you to one of their friend groups. Now imagine you have the opportunity to learn more about the people in each of those friend groups before deciding which one to get together with.    What are the most important identities, attributes, or characteristics you look for in a potential friend? (list up to 10)    Of those 10 you just listed, which 3 are the most important?    Now, of those 3 you just mentioned, which is the most important? | Imagine you are moving into a new apartment complex. Currently, there are multiple identical units available, and you have the chance to learn more about the people who would be living next door to each of those units.    What are the most important identities, attributes, or characteristics you would like to learn about and have in a potential neighbor? (list up to 10)    Of those 10 you just listed, which 3 are the most important?    Now, of those 3 you just mentioned, which is the most important? | Imagine you are looking for a new job and you just received two offers. These offers are very similar in terms of responsibilities, commutes, and salaries & benefits. The main difference between these positions would be the people with whom you are working. Now imagine you have the chance to learn more about the other employees at these two companies.    What are the most important identities, attributes, or characteristics you would like to have in a potential coworker? (list up to 10)    Of those 10 you just listed, which 3 are the most important?    Now, of those 3 you just mentioned, which is the most important? |

**Appendix B: Codebook**

**Directions**

I am interested in which identities are most important to people, both in terms of how they identify (ex. being a Christian) as well as what identities they care about in others (ex. wanting their friends to also be Christian). For this content analysis, we will be measuring which types of identities and traits are most frequently mentioned by respondents on a survey. This codebook details different types of identities and characteristics, and your job is to mark which responses fall under each category.

I administered a survey that asked respondents which identities were most important to them for themselves and the people in their relationships. Each respondent provided up to 10 identities or characteristics. The unit of analysis for this content analysis will be each trait, meaning each respondent can have between 1 and 10 traits for which you will be making coding decisions. The unit of analysis will be made apparent in the coding sheet by appearing as an individual row.

Check the box if the variable is present (i.e. you believe that response fits into that category) and leave the box empty if the variable is absent (i.e. you do not believe the response fits into that category). For example, if the response says “18,” you would check the box for age and leave the rest of the boxes unchecked. Each response should be coded for just one variable.

If there is a trait, and you leave boxes unchecked across all variables because you do not think the response fits into any category, check the box for “other” if you can think of a category it fits in that you think we should add. During practice coding, make a note if there is a category you think we could add that multiple responses under “other” could fit under. If you do not know how to code it, code it as “ambiguous.” Otherwise, check the box for “none” for nonsubstantive answers such as “don’t know”, “Ok”, etc.

When coding, do NOT leave anything blank or ask someone else’s opinion. All coding must be done individually without comparing answers until after.

**Variables**

Worry less about the tense or part of speech and more about the substance.

For each variable, the respondent can say the general category (ex. “age” or “hobby”) or a specific value of that category (ex. “18” or “dancing”)

**Age**: How old someone is. Either says “age” generally or mentions a specific age.

**Gender**: Either says “gender” or “sex” generally or mentions man, woman, male, female, nonbinary etc.

**Sexuality**: Either says “sexuality” or “sexual orientation” generally or mentions a specific one ex. “gay”, “bisexual”, etc.

**Race, ethnicity, or nationality**: Either says “race”, “ethnicity”, or “nationality” generally or mentions a specific one ex. “white”, “African American”, “Irish”, etc.

**Education level**: How much schooling one has received. Mentions being educated or level of education, not just being smart

**Religion**: Someone’s religion or religious or spiritual beliefs. Can be religion or spirituality. Mentions being religious or a specific religion such as Christianity, being Muslim, etc.

**Family role:** This refers to someone’s familial relationships. Could be mention of a specific role in family (ex. sister, grandfather, etc.) or family in general. Or something referring to it, such as “marital status.”

**Occupation**: mentions their job (ex. bank) or field of employment (ex. work in finance)

**Political views***:* this could include explicitly naming a party (ex. Republican, Democrat) or ideology (ex. conservative, liberal) or a political view about a policy that people (either citizens of the mass public or politicians) would vote on.

**Hobbies and interests**: Code if respondent mentions an activity, hobby, or interest such as dancing or being an artist. Include pet ownership under this category.

**Physical body:** Mentions a physical aspect of someone’s body. Following the only code one rule, something not covered by demographics. Talks about someone’s physical appearance such as their attractiveness or their health such as being overweight, being healthy, or having a diagnosis.

**Personality, behaviors, and other traits**: This includes traits mentioned about someone’s personal traits, whether they are characteristics or behaviors. This can include personality traits (extraversion, agreeableness, conscientiousness, emotional stability, or openness to experience), positive characteristics (ex. kind, honest), negative characteristics (ex. criminal record), or behaviors (ex. being quiet). These are descriptions of what the person is like or how the person acts. This includes mentions of valence terms such as “good” or “bad.” Only code if it is a characteristic or behavior that does not also fall under another category. (ex. “goes to church” would be religion, “plays sports” would be hobbies and interests, etc.).

**Ambiguous:** Seems to be an answer to the question (i.e. naming an identity, characteristic, or trait), but we cannot tell from the data provided enough information or context to know what variable it should be coded as. If something has multiple definitions that could fit into different categories, code as ambiguous. Code if one cell contains multiple answers (i.e. list) that fall into multiple categories.

**None**: The word is not an identity or a trait (e.g. people, doesnt, oriented).

**Other**: For practice coding, if an identity or trait does not fit into any of the categories, code it as other. You should only code it as other if it does not fit into any of the other categories. Only code it is other if it is a trait or identity, otherwise code it as none. This should be if you know what the category would be ex. We added “physical appearance” because of “looks”). If you do not know how it should be coded, that would be ambiguous. Make notes in the column to the right if there are any categories you think we should add.

**Tables**

**Table 1**

*Number of Respondents who Filled Out Each Version of the Survey*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **General** | **Specific** | **Friend** | **Neighbor** | **Coworker** |
| 42 | 50 | 43 | 50 | 48 |

Note: The table above shows the raw number of respondents who received each of five versions of the survey. Those in ‘General’ and ‘Specific’ were asked about the identities most important to them and which of those they shared with different people in their lives. The ‘Friend,’ ‘Neighbor’, and ‘Coworker’ survey versions asked about desired identities and characteristics in hypothetical relationships of those three types.

**Table 2**

*Intercoder Agreement Statistics*

|  |  |  |
| --- | --- | --- |
| **Percent Agreement** | **Cohen’s Kappa** | **Krippendorff’s Alpha (nominal)** |
| 90.5% | 0.787 | 0.786 |

Note: The above table shows intercoder agreement statistics for the content analysis of the open-ended survey responses as calculated by ReCal. Percent Agreement is the raw percentage of responses where the two coders agreed. The statistics for Cohen’s Kappa and Krippendorff’s Alpha provide chance-corrected agreement, which is why they are lower than the raw agreement.

**Table 3**

*Important Self-Identities and How Often They are Shared in Relationships*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Category** | **N** | **Family** | **Friend(s)** | **Acquaintance(s)** | **Disliked** |
| Personality, Behaviors, & Traits | 254 | 64.6% | 60.2% | 45.3% | 21.3% |
| Demographics | 44 | 70.5% | 47.7% | 43.2% | 38.6% |
| Hobbies & Interests | 20 | 55.0% | 35.0% | 30.0% | 5.0% |
| Physical Characteristics | 18 | 61.1% | 50.0% | 38.9% | 38.9% |
| Political Views | 1 | 0.0% | 0.0% | 0.0% | 0.0% |
| None of the Above | 40 | 60.0% | 52.5% | 45.0% | 25.0% |

Note: The table shows, for respondents who were asked about their own identities in the general and specific survey versions, which types of identities were most common and which they shared with the people in their lives. The column ‘Category’ lists the different categories of identities. The column ‘N’ provides the raw count of each type of identity in total. The percentages under family, friend(s), acquaintance(s), and disliked are the percentage of each type of identity that were marked as shared in each of those relationships. For example, the 64.6% under ‘Family’ means that people selected identities that were coded under personality, behaviors, and traits as being shared with family members 64.6% of the time.

**Table 4**

*Important Self-Identities and How Often They are Shared in Relationships (General Survey Version)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Self - General (N = 164)** | | | | |
| **Category** | **N** | **Family** | **Friend(s)** | **Acquaintance(s)** | **Disliked** |
| Personality, Behaviors, Traits | 105 | 61.9% | 52.4% | 40.0% | 26.7% |
| Demographics | 20 | 80.0% | 50.0% | 30.0% | 25.0% |
| Hobbies and Interests | 7 | 71.4% | 42.9% | 42.9% | 14.3% |
| Physical Characteristics | 8 | 62.5% | 50.0% | 37.5% | 37.5% |
| Political Views | 1 | 0.0% | 0.0% | 0.0% | 0.0% |
| None of the Above | 23 | 52.2% | 52.2% | 26.1% | 21.7% |

Note: This table follows the same format as Table 3 but looks at responses provided only from people who received the general survey version.

**Table 5**

*Important Self-Identities and How Often They are Shared in Relationships (Specific Survey Version)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Self - Specific (N = 213)** | | | | |
| **Category** | **N** | **Family** | **Friend(s)** | **Acquaintance(s)** | **Disliked** |
| Personality, Behaviors, Traits | 149 | 66.4% | 65.8% | 49.0% | 17.4% |
| Demographics | 24 | 62.5% | 45.8% | 54.2% | 50.0% |
| Hobbies and Interests | 13 | 46.2% | 30.8% | 23.1% | 0.0% |
| Physical Characteristics | 10 | 60.0% | 50.0% | 40.0% | 40.0% |
| Political Views | 0 | NA | NA | NA | NA |
| None of the Above | 17 | 70.6% | 52.9% | 70.6% | 29.4% |

Note: This table follows the same format as Table 3 but looks at responses provided only from people who received the specific survey version.

**Table 6**

*Top 10 Most Prioritized Identities Across Three Different Types of Hypothetical Relationships*

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Top 10** | | |
| **Category** | **Coworker**  **(N = 238)** | **Friend**  **(N = 216)** | **Neighbor**  **(N = 231)** |
| Personality, Behaviors, and Traits | 87.0% | 76.4% | 65.8% |
| Demographics | 2.5% | 5.1% | 15.2% |
| Hobbies and Interests | 0.4% | 4.2% | 5.6% |
| Physical Characteristics | 1.7% | 2.8% | 0.9% |
| Political Views | 0.8% | 1.4% | 0.4% |
| None of the Above | 7.6% | 10.2% | 12.1% |

Note: This table shows the results for the distribution of types of identities mentioned as important in coworker, friend, and neighbor relationships. The columns all add up to 100 percent and were calculated by taking the raw count for each category and dividing that by the number of responses for each type of relationship. For example, the percentages in the coworker category were all calculated by taking the raw count of each type of identity mentioned by those who received the coworker version of the survey and dividing that by 238.

**Table 7**

*Top 3 Most Prioritized Identities Across Three Different Types of Hypothetical Relationships*

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Top 3** | | |
| **Category** | **Coworker**  **(N = 139)** | **Friend**  **(N = 124)** | **Neighbor**  **(N = 134)** |
| Personality, Behaviors, and Traits | 89.2% | 81.5% | 73.9% |
| Demographics | 2.2% | 4.8% | 11.9% |
| Hobbies and Interests | 0.7% | 4.0% | 3.0% |
| Physical Characteristics | 0.0% | 0.0% | 0.0% |
| Political Views | 0.0% | 1.6% | 0.7% |
| None of the Above | 7.9% | 8.1% | 10.4% |

Note: This table is the same as Table 6, except it looks only at the identities marked as being in the top 3 most important for each type of relationship, whereas the previous table looked at the top 10.

**Table 8**

*Most Prioritized Identity Across Three Different Types of Hypothetical Relationships*

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Top 1** | | |
| **Category** | **Coworker**  **(N = 46)** | **Friend**  **(N = 43)** | **Neighbor**  **(N = 49)** |
| Personality, Behaviors, and Traits | 85.0% | 77.0% | 78.0% |
| Demographics | 2.0% | 2.0% | 12.0% |
| Hobbies and Interests | 2.0% | 5.0% | 2.0% |
| Physical Characteristics | 0.0% | 0.0% | 0.0% |
| Political Views | 0.0% | 2.0% | 0.0% |
| None of the Above | 11.0% | 14.0% | 8.0% |

Note: This table is the same as Tables 6 and 7, except it looks only at the identities marked as being the single most important identity across each type of relationship.

**Table 9**

*Most Common Words in Open-Ended Responses About One’s Own Identity*

|  |  |
| --- | --- |
| **Self (N = 377)** | |
| **word** | **n** |
| honest | 33 |
| love | 22 |
| fun | 21 |
| loyal | 16 |
| trust | 16 |
| care | 15 |
| smart | 15 |
| hardwork | 14 |
| empathy | 13 |
| friend | 13 |

Note: The above table shows the raw count of how often certain words appeared in the open-ended responses about one’s own identity in the general and specific survey versions. The table shows only the 10 most common words that appeared in responses. This is after data cleaning and stemming to combine words such as “honest,” “honesty,” and “honestly” into “honest.”

**Table 10**

*Most Common Words in Open-Ended Responses About Hypothetical Relationships*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Coworker**  **(N = 256)** | | **Friend**  **(N = 246)** | | **Neighbor**  **(N = 262)** | |
| **word** | **n** | **word** | **n** | **word** | **n** |
| hardwork | 15 | honest | 18 | friend | 25 |
| honest | 15 | loyal | 16 | nice | 12 |
| smart | 8 | friend | 12 | quiet | 10 |
| reliable | 7 | trust | 12 | person | 9 |
| trust | 7 | fun | 10 | fun; honest | 8 |

Note: The above table shows the same results as in Table 9 but for the other three versions of the survey.

**Figures**

**Figure 1**

*Results from Content Analysis about which Identities were Most Important to Respondents*

A graph showing the top ten identity

AI-generated content may be incorrect.

Note: Bars are raw counts. Each bar shows the count of each type of identity listed in the top 10 most important identities to oneself in the general and specific survey versions.

1. This number of 233 is after removing respondents that did not consent to the survey, did not complete the survey, or put nonsensical answers to the survey questions (such as random letters or names of celebrities) as well as after removing duplicate responses from the same respondent. In cases where the same respondent answered the survey multiple times, I kept their first response and deleted their later ones. Respondents that provided at least one response that answered the question being asked were kept in the dataset. [↑](#footnote-ref-1)
2. I intentionally chose not to provide examples of identities in order to avoid priming participants to think about particular categories of identities; I left it open to see what identities come to mind for people automatically, whether they be related to their race, diet, hobbies, family, gender, etc. This is also because Americans may conceptualize identities differently from the way scholars do so in academia. [↑](#footnote-ref-2)
3. Since I removed some responses that were incomplete, nonsensical, or duplicates, the randomization across groups was not exactly even. [↑](#footnote-ref-3)
4. According to Granovetter (1973), “the strength of a tie is a (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie" (p. 1361). [↑](#footnote-ref-4)